



## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: MELRI, JUDITH  
MUNNICH, ARNOLD
- (ii) TITLE OF INVENTION: SURVIVAL MOTOR NEURON (SMN) GENE: A GENE  
FOR SPINAL MUSCULAR ATROPHY
- (iii) NUMBER OF SEQUENCES: 65
- (iv) CORRESPONDENCE ADDRESS:  
(A) ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP  
(B) STREET: PO BOX 747  
(C) CITY: FALLS CHURCH  
(D) STATE: VA  
(E) COUNTRY: USA  
(F) ZIP: 22040-0747
- (v) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:  
(A) APPLICATION NUMBER: 08/545,196  
(B) FILING DATE: 19-OCT-1995  
(C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:  
(A) NAME: FARACI, C. J.  
(B) REGISTRATION NUMBER: 32,350  
(C) REFERENCE/DOCKET NUMBER: 2121-110P
- (ix) TELECOMMUNICATION INFORMATION:  
(A) TELEPHONE: (703) 205-8000  
(B) TELEFAX: (703) 205-8050

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 347 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AATTTTAA TTTTGTAG AGACAGGGTC TCATTATGTT SCCCAGGGTG GTGTCAAGCT

60

CCAGGTGCA AGTGATCCCC CTACCTCCGC CTCCCAAAGT TGTGGGATTG TAGGCATGAG

120

CCACTGCAAG AAAACCTTAA CTGCAGCCTA ATAATTGTTT TCTTTGGGAT AACTTTTAAA	180
GTACATTAAA AGACTATCAA CTTAATTTCT GATCATATTT TGTTGAATAA AATAAGTAAA	240
ATGTCTTGTG AACAAAATGC TTTTAAACAT CCATATAAAG CTATCTATAT ATAGCTATCT	300
ATGTCTATAT AGCTATTTTT TTTAACTTCC TTTTATTTTC CTTACAG	347

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 444 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GTAAGTCTGC CAGCATTATG AAAGTGAATC TACTTTTGT AAAACTTTAT GGTGTGTGGA	60
AAACAAATGT TTTTGAACAG TTAAAAAGTT CAGATGTTAA AAAGTTGAAA GGTAAATGTA	120
AAACAATCAA TATTAAAGAA TTTTGATGCC AAAACTATTA GATAAAAGGT TAATCTACAT	180
CCCTACTAGA ATTCTCATAC TTAAGTGGT GGTATGTGG AAGAAACATA CTTTCACAAT	240
AAAGAGCTTT AGGATATGAT GCCATTTTAT ATCACTAGTA GGCAGACCAG CAGACTTTTT	300
TTTATTGTGA TATGGGATAA CCTAGGCATA CTGCACTGTA CACTCTGACA TATGAAGTGC	360
TCTAGTCAAG TTTAACTGGT GTCCACAGAG GACATGGTTT AACTGGAATT CGTCAAGCCT	420
CTGGTTCTAA TTTCTCATTT GCAG	444

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 347 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AATTTTAAA TTTTGTAG AGACAGGGTC TCATTATGTT GCCCAGGGTG GTGTCAAGCT	60
CGAGGTCTCA AGTGATCCCC CTACCTCCGC CTCCCAAAGT TGTGGGATTG TAGGCATGAG	120
CCACTGCAAG AAAACCTTAA CTGCAGCCTA ATAATTGTTT TCTTTGGGAT AACTTTTAAA	180
GTACATTAAA AGACTATCAA CTTAATTTCT GATCATATTT TGTTGAATAA AATAAGTAAA	240

ATGTCTTGTG AACAAATGC TTTTAAACAT CCATATAAAG CTATCTATAT ATAGCTATCT 300  
 ATATCTATAT AGCTATTTTT TTAACTTCC TTTTATTTTC CTTACAG 347

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 444 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GTAAGTCTGC CAGCATTATG AAAGTGAATC TTA CTTTGTG AAAACTTTAT GGT TGTGGA 60  
 AAACAAATGT TTTTGAACAG TTA AAAAGTT CAGATGTTAG AAAGTTGAAA GGT TAATGTA 120  
 AAACAATCAA TATTAAAGAA TTTTGATGCC AAAACTATTA GATAAAAGGT TAATCTACAT 180  
 CCCTACTAGA ATTCTCATAC TTA ACTGGTT GGT TGTGTGG AAGAAACATA CTTTCACAAT 240  
 AAAGAGCTTT AGGATATGAT GCCATTTTAT ATCACTAGTA GGCAGACCAG CAGACTTTTT 300  
 TTTATTGTGA TATGGGATAA CCTAGGCATA CTGCACTGTA CACTCTGACA TATGAAGTGC 360  
 TCTAGTCAAG TTAACTGGT GTCCACAGAG GACATGGTTT AACTGGAATT CGTCAAGCCT 420  
 CTGGTTCTAA TTTCTCATTT GCAG 444

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AGACTATCAA CTTAATTTCT GATCA 25

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TAAGGAATGT GAGCACCTTC CTTC

24

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GTAATAACCA AATGCAATGT GAA

23

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CTACAACACC CTTCTCACAG

20

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 294 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Met Ala Met Ser Ser Gly Gly Ser Gly Gly Gly Val Pro Glu Gln Glu  
 1 5 10 15  
 Asp Ser Val Leu Phe Arg Arg Gly Thr Gly Gln Ser Asp Asp Ser Asp  
 20 25 30  
 Ile Trp Asp Asp Thr Ala Leu Ile Lys Ala Tyr Asp Lys Ala Val Ala  
 35 40 45  
 Ser Phe Lys His Ala Leu Lys Asn Gly Asp Ile Cys Glu Thr Ser Gly  
 50 55 60  
 Lys Pro Lys Thr Thr Pro Lys Arg Lys Pro Ala Lys Lys Asn Lys Ser  
 65 70 75 80  
 Gln Lys Lys Asn Thr Ala Ala Ser Leu Gln Gln Trp Lys Val Gly Asp  
 85 90 95  
 Lys Cys Ser Ala Ile Trp Ser Glu Asp Gly Cys Ile Tyr Pro Ala Thr  
 100 105 110  
 Ile Ala Ser Ile Asp Phe Lys Arg Glu Thr Cys Val Val Val Tyr Thr  
 115 120 125  
 Gly Tyr Gly Asn Arg Glu Glu Gln Asn Leu Ser Asp Leu Leu Ser Pro  
 130 135 140  
 Ile Cys Glu Val Ala Asn Asn Ile Glu Gln Asn Ala Gln Glu Asn Glu  
 145 150 155 160  
 Asn Glu Ser Gln Val Ser Thr Asp Glu Ser Glu Asn Ser Arg Ser Pro  
 165 170 175  
 Gly Asn Lys Ser Asp Asn Ile Lys Pro Lys Ser Ala Pro Trp Asn Ser  
 180 185 190  
 Phe Leu Pro Pro Pro Pro Pro Met Pro Gly Pro Arg Leu Gly Pro Gly  
 195 200 205  
 Lys Pro Gly Leu Lys Phe Asn Gly Pro Pro Pro Pro Pro Pro Pro  
 210 215 220  
 Pro Pro His Leu Leu Ser Cys Trp Leu Pro Pro Phe Pro Ser Gly Pro  
 225 230 235 240  
 Pro Ile Ile Pro Pro Pro Pro Pro Ile Cys Pro Asp Ser Leu Asp Asp  
 245 250 255  
 Ala Asp Ala Leu Gly Ser Met Leu Ile Ser Trp Tyr Met Ser Gly Tyr  
 260 265 270  
 His Thr Gly Tyr Tyr Met Gly Phe Arg Gln Asn Gln Lys Glu Gly Arg  
 275 280 285  
 Cys Ser His Ser Leu Asn  
 290

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1582 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii). MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CGGGGCCCCA CGCTGCGCAC CCGCGGGTTT GCTATGGCGA TGAGCAGCGG CGGCAGTGGT	60
GGCGGCGTCC CGGAGCAGGA GGATTCCGTG CTGTTCCGGC GCGGCACAGG CCAGAGCGAT	120
GATTCTGACA TTTGGGATGA TACAGCACTG ATAAAAGCAT ATGATAAAGC TGTGGCTTCA	180
TTTAAGCATG CTCTAAAGAA TGGTGACATT TGTGAAACTT CGGGTAAACC AAAAACCACA	240
CCTAAAAGAA AACCTGCTAA GAAGAATAAA AGCCAAAAGA AGAATACTGC AGCTTCCTTA	300
CAACAGTGGA AAGTTGGGGA CAAATGTTCT GCCATTTGGT CAGAAGACGG TTGCATTTAC	360
CCAGCTACCA TTGCTTCAAT TGATTTTAAG AGAGAAACCT GTGTTGTGGT TTACTACTGGA	420
TATGGAAATA GAGAGGAGCA AAATCTGTCC GATCTACTTT CCCCAATCTG TGAAGTAGCT	480
AATAATATAG AACAGAATGC TCAAGAGAAT GAAAATGAAA GCCAAGTTTC AACAGATGAA	540
AGTGAGAACT CCAGGTCTCC TGGAAATAAA TCAGATAACA TCAAGCCCAA ATCTGCTCCA	600
TGGAACCCCT TTCTCCCTCC ACCACCCCCC ATGCCAGGGC CAAGACTGGG ACCAGGAAAG	660
CCAGGTCTAA AATTCAATGG CCCACCACCG CCACCGCCAC CACCACCACC CCACTTACTA	720
TCATGCTGGC TGCCTCCATT TCCTTCTGGA CCACCAATAA TTCCCCCACC ACCTCCCATA	780
TGTCCAGATT CTCTTGATGA TGCTGATGCT TTGGGAAGTA TGTTAATTTT ATGGTACATG	840
AGTGGCTATC ATACTGGCTA TTATATGGGT TTTAGACAAA ATCAAAAAGA AGGAAGGTGC	900
TCACATTCCT TAAATTAAGG AGAAATGCTG GCATAGAGCA GCACTAAATG ACACCACTAA	960
AGAAACGATC AGACAGATCT GGAATGTGAA GCGTTATAGA AGATAACTGG CCTCATTTCT	1020
TCAAAATATC AAGTGTGGG AAAGAAAAAA GGAAGTGGAA TGGGTAATCT TTCTTGATTA	1080
AAAGTTATGT AATAACCAA TGCAATGTGA AATATTTTAC TGGACTCTTT TGAAAAACCA	1140
TCTGTAAAAG ACTGAGGTGG GGGTGGGAGG CCAGCACGGT GGTGAGGCAG TTGAGAAAAT	1200
TTGAATGTGG ATTAGATTTT GAATGATATT GGATAATTAT TGGTAATTTT ATGGCCTGTG	1260
AGAAGGGTGT TGAGTTTAT AAAAGACTGT CTTAATTTGC ATACTTAAGC ATTTAGGAAT	1320
GAAGTGTTAG AGTGCTTAA AATGTTTCAA ATGGTTTAAC AAAATGTATG TGAGGCGTAT	1380
GTGGCAAAAT GTTACAGAAT CTAAGTGGTG GACATGGCTG TTCATTGTAC TGTTTTTTTC	1440
TATCTTCTAT ATGTTTAAAA GTATATAATA AAAATATTTA ATTTTTTTTT AAAAAAAAAA	1500

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 1560  
 AAAAAAAAAA AAAAAAAAAA AA 1582

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1408 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

AATTTTAA TTTTGTAG AGACAGGGTC TCATTATGTT GCCCAGGGTG GTGTCAAGCT 60  
 CCAGGTCTCA AGTGATCCCC CTACCTCCGC CTCCCAAAGT TGTGGGATTG TAGGCATGAG 120  
 CCACTGCAAG AAAACCTTAA CTGCAGCCTA ATAATTGTTT TCTTTGGGAT AACTTTTAAA 180  
 GTACATTAAA AGACTATCAA CTTAATTTCT GATCATATTT TGTTGAATAA AATAAGTAAA 240  
 ATGTCTTG TG AACAAATGC TTTTAAACAT CCATATAAAG CTATCTATAT ATAGCTATCT 300  
 ATATCTATAT AGCTATTTTT TTTAACTTCC TTTTATTTTC CTTACAGGGT TTTAGACAAA 360  
 ATCAAAAAGA AGGAAGGTGC TCACATTCCT TAAATTAAGG AGTAAGTCTG CCAGCATTAT 420  
 GAAAGTGAAT CTTACTTTTG TAAAACTTTA TGGTTTGTGG AAAACAAATG TTTTGAACA 480  
 GTTAAAAAGT TCAGATGTTA GAAAGTTGAA AGGTAAATGT AAAACAATCA ATATTAAAGA 540  
 ATTTTGATGC CAAAACTATT AGATAAAAGG TTAATCTACA TCCCTACTAG AATTCTCATA 600  
 CTTAACTGGT TGGTTGTGTG GAAGAAACAT ACTTTCACAA TAAAGAGCTT TAGGATATGA 660  
 TGCCATTTTA TATCACTAGT AGGCAGACCA GCAGACTTTT TTTTATTGTG ATATGGGATA 720  
 ACCTAGGCAT ACTGCACTGT ACACTCTGAC ATATGAAGTG CTCTAGTCAA GTTTAACTGG 780  
 TGTCCACAGA GGACATGGTT TAACTGGAAT TCGTCAAGCC TCTGGTTCTA ATTTCTCATT 840  
 TGCAGGAAAT GCTGGCATAG AGCAGCACTA AATGACACCA CTAAAGAAAC GATCAGACAG 900  
 ATCTGGAATG TGAAGCGTTA TAGAAGATAA CTGGCCTCAT TTCTTCAAAA TATCAAGTGT 960  
 TGGGAAAGAA AAAAGGAAGT GGAATGGGTA ACTCTTCTTG ATTAAAAGTT ATGTAATAAC 1020  
 CAAATGCAAT GTGAAATATT TTA CTGGACT CTTTGGAAAA ACCATCTGTA AAAGACTGAG 1080  
 GTGGGGGTGG GAGGCCAGCA CGGTGGTGAG GCAGTTGAGA AAATTTGAAT GTGGATTAGA 1140  
 TTTTGAATGA TATTGGATAA TTATTGGTAA TTTTATGGCC TGTGAGAAGG GTGTTGTAGT 1200  
 TTATAAAAGA CTGTCTTAAT TTGCATACTT AAGCATTTAG GAATGAAGTG TTAGAGTGTC 1260

TTAAATGTT TCAAATGGTT TAACAAAATG TATGTGAGGC GTATGTGGCA AAATGTTACA 1320  
 GAATCTAACT GGTGGACATG GCTGTTTATT GTACTGTTTT TTTCTATCTT CTATATGTTT 1380  
 AAAAGTATAT AATAAAAATA TTTAATTT 1408

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1582 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CGGGGCCCCA CGCTGCGCAT CCGCGGGTTT GCTATGGCGA TGAGCAGCGG CGGCAGTGGT 60  
 GGCGGCGTCC CGGAGCAGGA GGATTCCGTG CTGTTCCGGC GCGGCACAGG CCAGAGCGAT 120  
 GATTCTGACA TTTGGGATGA TACAGCACTG ATAAAAGCAT ATGATAAAGC TGTGGCTTCA 180  
 TTTAAGCATG CTCTAAAGAA TGGTGACATT TGTGAACTT CGGGTAAACC AAAAACCACA 240  
 CCTAAAAGAA AACCTGCTAA GAAGAATAAA AGCCAAAAGA AGAATACTGC AGCTTCCTTA 300  
 CAACAGTGGA AAGTTGGGGA CAAATGTTCT GCCATTTGGT CAGAAGACGG TTGCATTTAC 360  
 CCAGCTACCA TTGCTTCAAT TGATTTTAAG AGAGAAACCT GTGTTGTGGT TTACACTGGA 420  
 TATGGAAATA GAGAGGAGCA AAATCTGTCC GATCTACTTT CCCCAATCTG TGAAGTAGCT 480  
 AATAATATAG AACAGAATGC TCAAGAGAAT GAAAATGAAA GCCAAGTTTC AACAGATGAA 540  
 AGTGAGAACT CCAGGTCTCC TGGAAATAAA TCAGATAACA TCAAGCCCAA ATCTGCTCCA 600  
 TGGAATCTT TTCTCCCTCC ACCACCCCCC ATGCCAGGGC CAAGACTGGG ACCAGGAAAG 660  
 CCAGGTCTAA AATTCAATGG CCCACCACCG CCACCGCCAC CACCACCACC CCACTTACTA 720  
 TCATGCTGGC TGCCTCCATT TCCTTCTGGA CCACCAATAA TTCCCCCACC ACCTCCCATA 780  
 TGTCCAGATT CTCTTGATGA TGCTGATGCT TTGGGAAGTA TGTTAATTTT ATGGTACATG 840  
 AGTGGCTATC ATACTGGCTA TTATATGGGT TTCAGACAAA ATCAAAAAGA AGGAAGGTGC 900  
 TCACATTCTT TAAATTAAGG AGAAATGCTG GCATAGAGCA GCACTAAATG ACACCACTAA 960  
 AGAAACGATC AGACAGATCT GGAATGTGAA GCGTTATAGA AGATAACTGG CCTCATTTCT 1020  
 TCAAAATATC AAGTGTTGGG AAAGAAAAAA GGAAGTGGAA TGGGTAATCT TTCTTGATTA 1080  
 AAAGTTATGT AATAACCAA TGCAATGTGA AATATTTTAC TGGACTCTTT TGAAAAACCA 1140  
 TCTGTAAAAG ACTGGGGTGG GGGTGGGAGG CCAGCACGGT GGTGAGGCAG TTGAGAAAAT 1200



TTGAATGTGG ATTAGATTTT GAATGATATT GGATAATTAT TGGTAATTTT ATGGCCTGTG	1260
AGAAGGGTGT TGTAGTTTAT AAAAGACTGT CTTAATTTGC ATACTTAAGC ATTTAGGAAT	1320
GAAGTGTTAG AGTGTCTTAA AATGTTTCAA ATGGTTTAAC AAAATGTATG TGAGGCGTAT	1380
GTGGCAAAAT GTTACAGAAT CTAACGGTG GACATGGCTG TTCATTGTAG TGTTTTTTTC	1440
TATCTTCTAT ATGTTTAAAA GTATATAATA AAAATATTTA ATTTTTTTTT AAAAAAAAAA	1500
AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA	1560
AAAAAAAAAA AAAAAAAAAA AA	1582

## (2) INFORMATION FOR SEQ ID NO:13:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1408 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AATTTTAAAA TTTTGTAG AGACAGGGTC TCATTATGTT GCCCAGGGTG GTGTCAAGCT	60
CCAGGTCTCA AGTGATCCCC CTACCTCCGC CTCCCAAAGT TGTGGGATTG TAGGCATGAG	120
CCACTGCAAG AAAACCTTAA CTGCAGCCTA ATAATTGTTT TCTTTGGGAT AACTTTTAAA	180
GTACATTAAA AGACTATCAA CTTAATTTCT GATCATATTT TGTTGAATAA AATAAGTAAA	240
ATGTCTTG TG AACAAAATGC TTTTAAACAT CCATATAAAG CTATCTATAT ATAGCTATCT	300
ATGTCTATAT AGCTATTTTT TTTAACTTCC TTTTATTTTC CTTACAGGGT TTCAGACAAA	360
ATCAAAAAGA AGGAAGGTGC TCACATTCCT TAAATTAAGG AGTAAGTCTG CCAGCATTAT	420
GAAAGTGAAT CTTACTTTTG TAAACTTTA TGGTTTGTGG AAAACAAATG TTTTGAACA	480
GTAAAAAGT TCAGATGTTA AAAAGTTGAA AGGTTAATGT AAAACAATCA ATATTAAAGA	540
ATTTTGATGC CAAACTATT AGATAAAAGG TTAATCTACA TCCCTACTAG AATTCTCATA	600
CTTAAGTGGT TGGTTATGTG GAAGAAACAT ACTTTCACAA TAAAGAGCTT TAGGATATGA	660
TGCCATTTTA TATCACTAGT AGGCAGACCA GCAGACTTTT TTTTATTGTG ATATGGGATA	720
ACCTAGGCAT ACTGCACTGT ACACTCTGAC ATATGAAGTG CTCTAGTCAA GTTAACTGG	780
TGTCCACAGA GGACATGGTT TAACTGGAAT TCGTCAAGCC TCTGGTTCTA ATTTCTCATT	840
TGCAGGAAAT GCTGGCATAG AGCAGCACTA AATGACACCA CTAAAGAAAC GATCAGACAG	900
ATCTGGAATG TGAAGCGTTA TAGAAGATAA CTGGCCTCAT TTCTTCAAAA TATCAAGTGT	960

TGGGAAAGAA AAAAGGAAGT GGAATGGGTA ACTCTTCTTG ATTAAGAGTT ATGTAATAAC	1020
CAAATGCAAT GTGAAATATT TTACTGGACT CTTTGGAAAA ACCATCTGTA AAAGACTGGG	1080
GTGGGGGTGG GAGGCCAGCA CGGTGGTGAG GCAGTTGAGA AAATTTGAAT GTGGATTAGA	1140
TTTTGAATGA TATTGGATAA TTATTGGTAA TTTTATGGCC TGTGAGAAGG GTGTTGTAGT	1200
TTATAAAGA CTGTCTTAAT TTGCATACTT AAGCATTTAG GAATGAAGTG TTAGAGTGTC	1260
TTAAATGTT TCAAATGGTT TAACAAAATG TATGTGAGGC GTATGTGGCA AAATGTTACA	1320
GAATCTAACT GGTGGACATG GCTGTTTCATT GTACTGTTTT TTTCTATCTT CTATATGTTT	1380
AAAAGTATAT AATAAAAATA TTTAATTT	1408

## (2) INFORMATION FOR SEQ ID NO:14:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 372 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

ACCTGANCCC AGANGGTCAAG GCTGCAGTG AGACGAGATT GCNCCACTGC CCTCCACCCT	60
GGGTGATAAG AGTGGGACCC TGTNTCAAAA CACACACACA CACACACACA CACACACACA	120
CACACACACA CACACTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCTCTCA	180
AAAACACTTG GTCTGTTATT TTTNCGAAAT TGTCAGTCAT AGTTATCTGT TAGACCAAAG	240
CTGNGTAAGN ACATTTATTA CATTGCCTCC TACAACTTCA TCAGCTAATG TATTTGCTAT	300
ATAGCAATTA CATATNGGNA TATATTATCT TNAGGGGATG GCCANGTNAT AAAACTGTCA	360
CTGAGGAAAG GA	372

## (2) INFORMATION FOR SEQ ID NO:15:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 294 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CCTCCACCT NAGCCTCCCC AGTAGCTAGG ACTATAGGCG TGCNCCACCA AGCTCAGCTA	60
TTTTTNNTAT TTAGTAGAGA CGGGGTTTCG GCANGCTTAG GCCTCGTNTC GAACTCCAGT	120
GTGTGTGTGT GTGTGTGTGT GTGTGTGTGT GTGTGTGTGT GTGTGTGTGT AGATATTTAT	180
TCCCCCTCCC CTTGGAAAA GTAAGTAAGC TCCTACTAGG AATTAAAAAC CTGCTTGATC	240
TATATAAGA CAAACAAGGA AAGACAAACA TGGGGGCAGG AAGGAAGGCA GATC	294

## (2) INFORMATION FOR SEQ ID NO:16:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 141 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

TCGAGGTAGA TTTGTATTAT ATCCCATGTA CACACACACA CACACACACA CACACACACA	60
CACACACAGA CTTAATCTGT TTACAGAAAT AAAAGGAATA AAATACCGTT TCTACTATAC	120
ACCAAAACTA GCCATCTTGA C	141

## (2) INFORMATION FOR SEQ ID NO:17:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 305 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CCCTGAGAAG GCTTCCTCCT GAGTATGCAT AAACATTAC AGCTTGCATG CGTGTGTGTG	60
TGTGTGTGTG TGTGTATGTT TGCTTGCACT GTAAAAACAA TTGCAACATC AACAGAAATA	120
AAAATTAAAG GAATAATTCT CCTCCGACTC TGCCGTTCCA TCCAGTGAAA CTCTTCATTC	180
TGGGGTAAAG TTCCTTCAGT TCTTTCATAG ATAGGTATAT ACTTCATAAG TCAAACAATC	240
AGGCTGGGTG CAGTAGCTCA TGCCTGTAAT CCCAGCCCTT TGGGAGGCCG AGCTGGGCAG	300

ATCGA

305

## (2) INFORMATION FOR SEQ ID NO:18:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 350 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```

TCCACCCGCC TTGGCCTCCC AAAGCNCCTGG GATTACAGGC GTGACTGCCG CACCCAGCTG      60
TAAACTGGNT TNNTAATGGT AGATTTTNAG GTATTAACAA TAGATAAAAA GATACTTTTN      120
GGCATACTGT GTATTGGGAT GGGGTTAGAA CAGGTGTNCT ACCCAAGACA TTTACTTAAA      180
ATCGCCCTCG AAATGCTATG TGAGCTGTGT GTGTGTGTGT GTGTGTGTGT GTATTAAGGA      240
AAAGCATGAA AGTATTTATG CTTGATTTTT TTTTNNACT CATAGCTTCA TAGTGGANCA      300
GATACATAGT CTAAATCAAA ATGTTTAAAC TTTTATGTC ACTTGCTGTC      350

```

## (2) INFORMATION FOR SEQ ID NO:19:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 278 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

```

Met Ala Met Ser Ser Gly Gly Ser Gly Gly Gly Val Pro Glu Gln Glu
1           5           10           15
Asp Ser Val Leu Phe Arg Arg Gly Thr Gly Gln Ser Asp Asp Ser Asp
20        25        30
Ile Trp Asp Asp Thr Ala Leu Ile Lys Ala Tyr Asp Lys Ala Val Ala
35        40        45
Ser Phe Lys His Ala Leu Lys Asn Gly Asp Ile Cys Glu Thr Ser Gly
50        55        60
Lys Pro Lys Thr Thr Pro Lys Arg Lys Pro Ala Lys Lys Asn Lys Ser
65        70        75        80
Gln Lys Lys Asn Thr Ala Ala Ser Leu Gln Gln Trp Lys Val Gly Asp

```

85	90	95
Lys Cys Ser Ala Ile Trp Ser Glu Asp Gly Cys Ile Tyr Pro Ala Thr		
100	105	110
Ile Ala Ser Ile Asp Phe Lys Arg Glu Thr Cys Val Val Val Tyr Thr		
115	120	125
Gly Tyr Gly Asn Arg Glu Glu Gln Asn Leu Ser Asp Leu Leu Ser Pro		
130	135	140
Ile Cys Glu Val Ala Asn Asn Ile Glu Gln Asn Ala Gln Glu Asn Glu		
145	150	155
Asn Glu Ser Gln Val Ser Thr Asp Glu Ser Glu Asn Ser Arg Ser Pro		
165	170	175
Gly Asn Lys Ser Asp Asn Ile Lys Pro Lys Ser Ala Pro Trp Asn Ser		
180	185	190
Phe Leu Pro Pro Pro Pro Pro Met Pro Gly Pro Arg Leu Gly Pro Gly		
195	200	205
Lys Pro Gly Leu Lys Phe Asn Gly Pro Pro Pro Pro Pro Pro Pro Pro		
210	215	220
Pro Pro His Leu Leu Ser Cys Trp Leu Pro Pro Phe Pro Ser Gly Pro		
225	230	235
Pro Ile Ile Pro Pro Pro Pro Pro Ile Cys Pro Asp Ser Leu Asp Asp		
245	250	255
Ala Asp Ala Leu Gly Ser Met Leu Ile Ser Trp Tyr Met Ser Gly Tyr		
260	265	270
His Thr Gly Tyr Tyr Met		
275		

## (2) INFORMATION FOR SEQ ID NO:20:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 885 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 18..881

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CGGCGTGGTA GCAGGCC ATG GCG ATG GGC AGT GGC GGA GCG GGC TCC GAG	50
Met Ala Met Gly Ser Gly Gly Ala Gly Ser Glu	
1 5 10	
CAG GAA GAT ACG GTG CTG TTC CGG CGT GGC ACC GGC CAG AGT GAT GAT	98

Gln Glu Asp Thr Val Leu Phe Arg Arg Gly Thr Gly Gln Ser Asp Asp	15	20	25	
TCT GAC ATT TGG GAT GAT ACA GCA TTG ATA AAA GCT TAT GAT AAA GCT				146
Ser Asp Ile Trp Asp Asp Thr Ala Leu Ile Lys Ala Tyr Asp Lys Ala	30	35	40	
GTG GCT TCC TTT AAG CAT GCT CTA AAG AAC GGT GAC ATT TGT GAA ACT				194
Val Ala Ser Phe Lys His Ala Leu Lys Asn Gly Asp Ile Cys Glu Thr	45	50	55	
CCA GAT AAG CCA AAA GGC ACA GCC AGA AGA AAA CCT GCC AAG AAG AAT				242
Pro Asp Lys Pro Lys Gly Thr Ala Arg Arg Lys Pro Ala Lys Lys Asn	60	65	70	75
AAA AGC CAA AAG AAG AAT GCC ACA ACT CCC TTG AAA CAG TGG AAA GTT				290
Lys Ser Gln Lys Lys Asn Ala Thr Thr Pro Leu Lys Gln Trp Lys Val	80	85	90	
GGT GAC AAG TGT TCT GCT GTT TGG TCA GAA GAC GGC TGC ATT TAC CCA				338
Gly Asp Lys Cys Ser Ala Val Trp Ser Glu Asp Gly Cys Ile Tyr Pro	95	100	105	
GCT ACT ATT ACG TCC ATT GAC TTT AAG AGA GAA ACC TGT GTC GTG GTT				386
Ala Thr Ile Thr Ser Ile Asp Phe Lys Arg Glu Thr Cys Val Val Val	110	115	120	
TAT ACT GGA TAT GGA AAC AGA GAG GAG CAA AAC TTA TCT GAC CTA CTT				434
Tyr Thr Gly Tyr Gly Asn Arg Glu Glu Gln Asn Leu Ser Asp Leu Leu	125	130	135	
TCC CCG ACC TGT GAA GTA GCT AAT AGT ACA GAA CAG AAC ACT CAG GAG				482
Ser Pro Thr Cys Glu Val Ala Asn Ser Thr Glu Gln Asn Thr Gln Glu	140	145	150	155
AAT GAA AGT CAA GTT TCC ACA GAC GAC AGT GAA CAC TCC TCC AGA TCG				530
Asn Glu Ser Gln Val Ser Thr Asp Asp Ser Glu His Ser Ser Arg Ser	160	165	170	
CTC AGA AGT AAA GCA CAC AGC AAG TCC AAA GCT GCT CCG TGG ACC TCA				578
Leu Arg Ser Lys Ala His Ser Lys Ser Lys Ala Ala Pro Trp Thr Ser	175	180	185	
TTT CTT CCT CCA CCA CCC CCA ATG CCA GGG TCA GGA TTA GGA CCA GGA				626
Phe Leu Pro Pro Pro Pro Met Pro Gly Ser Gly Leu Gly Pro Gly	190	195	200	
AAG CCA GGT CTA AAA TTC AAC GGC CCG CCG CCG CCG CCT CCA CTA CCC				674
Lys Pro Gly Leu Lys Phe Asn Gly Pro Pro Pro Pro Pro Pro Leu Pro	205	210	215	
CCT CCC CCC TTC CTG CCG TGC TGG ATG CCC CCG TTC CCT TCA GGA CCA				722
Pro Pro Pro Phe Leu Pro Cys Trp Met Pro Pro Phe Pro Ser Gly Pro	220	225	230	235
CCA ATA ATC CCG CCA CCC CCT CCC ATC TCT CCC GAC TGT CTG GAT GAC				770
Pro Ile Ile Pro Pro Pro Pro Pro Ile Ser Pro Asp Cys Leu Asp Asp	240	245	250	
ACT GAT GCC CTG GGC AGT ATG CTA ATC TCT TGG TAC ATG AGT GGC TAC				818
Thr Asp Ala Leu Gly Ser Met Leu Ile Ser Trp Tyr Met Ser Gly Tyr				

255	260	265	
CAC ACT GGC TAC TAT ATG GGT TTC AGA CAA AAT AAA AAA GAA GGA AAG			866
His Thr Gly Tyr Tyr Met Gly Phe Arg Gln Asn Lys Lys Glu Gly Lys			
270	275	280	
TGC TCA CAT ACA AAT TAAG			885
Cys Ser His Thr Asn			
285			

## (2) INFORMATION FOR SEQ ID NO:21:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 288 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Met	Ala	Met	Gly	Ser	Gly	Gly	Ala	Gly	Ser	Glu	Gln	Glu	Asp	Thr	Val	
1				5					10					15		
Leu	Phe	Arg	Arg	Gly	Thr	Gly	Gln	Ser	Asp	Asp	Ser	Asp	Ile	Trp	Asp	
			20					25					30			
Asp	Thr	Ala	Leu	Ile	Lys	Ala	Tyr	Asp	Lys	Ala	Val	Ala	Ser	Phe	Lys	
		35				40					45					
His	Ala	Leu	Lys	Asn	Gly	Asp	Ile	Cys	Glu	Thr	Pro	Asp	Lys	Pro	Lys	
	50					55					60					
Gly	Thr	Ala	Arg	Arg	Lys	Pro	Ala	Lys	Lys	Asn	Lys	Ser	Gln	Lys	Lys	
	65				70					75				80		
Asn	Ala	Thr	Thr	Pro	Leu	Lys	Gln	Trp	Lys	Val	Gly	Asp	Lys	Cys	Ser	
					85				90					95		
Ala	Val	Trp	Ser	Glu	Asp	Gly	Cys	Ile	Tyr	Pro	Ala	Thr	Ile	Thr	Ser	
			100					105					110			
Ile	Asp	Phe	Lys	Arg	Glu	Thr	Cys	Val	Val	Val	Tyr	Thr	Gly	Tyr	Gly	
		115					120						125			
Asn	Arg	Glu	Glu	Gln	Asn	Leu	Ser	Asp	Leu	Leu	Ser	Pro	Thr	Cys	Glu	
	130					135					140					
Val	Ala	Asn	Ser	Thr	Glu	Gln	Asn	Thr	Gln	Glu	Asn	Glu	Ser	Gln	Val	
	145					150				155				160		
Ser	Thr	Asp	Asp	Ser	Glu	His	Ser	Ser	Arg	Ser	Leu	Arg	Ser	Lys	Ala	
				165					170					175		
His	Ser	Lys	Ser	Lys	Ala	Ala	Pro	Trp	Thr	Ser	Phe	Leu	Pro	Pro	Pro	
				180				185					190			
Pro	Pro	Met	Pro	Gly	Ser	Gly	Leu	Gly	Pro	Gly	Lys	Pro	Gly	Leu	Lys	
		195					200					205				

Phe Asn Gly Pro Pro Pro Pro Pro Pro Leu Pro Pro Pro Phe Leu  
 210 215 220  
 Pro Cys Trp Met Pro Pro Phe Pro Ser Gly Pro Pro Ile Ile Pro Pro  
 225 230 235 240  
 Pro Pro Pro Ile Ser Pro Asp Cys Leu Asp Asp Thr Asp Ala Leu Gly  
 245 250 255  
 Ser Met Leu Ile Ser Trp Tyr Met Ser Gly Tyr His Thr Gly Tyr Tyr  
 260 265 270  
 Met Gly Phe Arg Gln Asn Lys Lys Glu Gly Lys Cys Ser His Thr Asn  
 275 280 285

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3271 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 104..184

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 364..435

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 637..756

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 921..1121

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1265..1417

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1605..1700

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1810..1920

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2214..2261

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:



CCTCCCGGGC ACCGTACTGT TCCGCTCCCA GAAGCCCCGG GCGCCGGAAG TCGTCACTCT	60
TAAGAAGGGA CGGGGCCCCA CGCTGCGCAC CCGCGGGTTT GCT ATG GCG ATG AGC	115
Met Ala Met Ser	
1	
AGC GGC GGC AGT GGT GGC GGC GTC CCG GAG CAG GAG GAT TCC GTG CTG	163
Ser Gly Gly Ser Gly Gly Gly Val Pro Glu Gln Glu Asp Ser Val Leu	
5 10 15 20	
TTC CGG CGC GGC ACA GGC CAG GTGAGGTCGC AGCCAGTGCA GTCTCCCTAT	214
Phe Arg Arg Gly Thr Gly Gln	
25	
TAGCGCTCTC AGCACCTTC TTCCGGCCCCA ACTCTCCTTC CGCAGTGTA TTTTGTTATG	274
TGTGGATTAA GATGACTCTT GGTACTAACA TACATTTTCT GATTAAACCT ATCTGNACAT	334
GAGTTGTTTT TATTTCTTAC CCTTTCCAG AGC GAT GAT TCT GAC ATT TGG GAT	387
Ser Asp Asp Ser Asp Ile Trp Asp	
30 35	
GAT ACA GCA CTG ATA AAA GCA TAT GAT AAA GCT GTG GCT TCA TTT AAG	435
Asp Thr Ala Leu Ile Lys Ala Tyr Asp Lys Ala Val Ala Ser Phe Lys	
40 45 50	
GTATGAAATG CTTGNNTAGT CGTTTTCTTA TTTTCTCGTT ATTCATTTGG AAAGGAATTG	495
ATAACATACG ATAAAGTGTT AAAGGTGCTT TCTGAGGTGA CGGAGCCTTG AGACTAGCTT	555
ATAGTAGTAA CTGGGTTATG TCGTGACTTT TATTCTGTGC ACCACCCTGT AACATGTACA	615
TTTTTATTCC TATTTTCGTA G CAT GCT CTA AAG AAT GGT GAC ATT TGT GAA	666
His Ala Leu Lys Asn Gly Asp Ile Cys Glu	
55 60	
ACT TCG GGT AAA CCA AAA ACC ACA CCT AAA AGA AAA CCT GCT AAG AAG	714
Thr Ser Gly Lys Pro Lys Thr Thr Pro Lys Arg Lys Pro Ala Lys Lys	
65 70 75	
AAT AAA AGC CAA AAG AAG AAT ACT GCA GCT TCC TTA CAA CAG	756
Asn Lys Ser Gln Lys Lys Asn Thr Ala Ala Ser Leu Gln Gln	
80 85 90	
GTTATTTTAA AATGTTGAGG ATTTAACTTC AAAGGATGTC TCATTAGTCC TTATTTAATA	816
GTGTAAAATG TCTTTAACTG CCTGCAGGTC GATCAAAACG AGATGATAGT TTGCCCTCTT	876
CAAAAGAAAT GTGTGCATGT ATATATCTTT GATTTCTTTT GTAG TGG AAA GTT GGG	932
Trp Lys Val Gly	
95	
GAC AAA TGT TCT GCC ATT TGG TCA GAA GAC GGT TGC ATT TAC CCA GCT	980
Asp Lys Cys Ser Ala Ile Trp Ser Glu Asp Gly Cys Ile Tyr Pro Ala	
100 105 110	
ACC ATT GCT TCA ATT GAT TTT AAG AGA GAA ACC TGT GTT GTG GTT TAC	1028
Thr Ile Ala Ser Ile Asp Phe Lys Arg Glu Thr Cys Val Val Val Tyr	
115 120 125	

ACT GGA TAT GGA AAT AGA GAG GAG CAA AAT CTG TCC GAT CTA CTT TCC 1076  
 Thr Gly Tyr Gly Asn Arg Glu Glu Gln Asn Leu Ser Asp Leu Leu Ser  
 130 135 140

CCA ATC TGT GAA GTA GCT AAT AAT ATA GAA CAG AAT GCT CAA GAG 1121  
 Pro Ile Cys Glu Val Ala Asn Asn Ile Glu Gln Asn Ala Gln Glu  
 145 150 155

GTAAGGATAC AAAAAAAAAA AAATTCAATT TCTGGAAGCA GAGACTAGAT GAGAAACTGT 1181

TAAACAGTAT ACACCACCGA GGCATTAATT TTTTCTTAAT CACACCCTTA TAACAAAAAC 1241

CTGCATATTT TTTCTTTTAA AAG AAT GAA AAT GAA AGC CAA GTT TCA ACA GAT 1294  
 Asn Glu Asn Glu Ser Gln Val Ser Thr Asp  
 160 165

GAA AGT GAG AAC TCC AGG TCT CCT GGA AAT AAA TCA GAT AAC ATC AAG 1342  
 Glu Ser Glu Asn Ser Arg Ser Pro Gly Asn Lys Ser Asp Asn Ile Lys  
 170 175 180

CCC AAA TCT GCT CCA TGG AAC TCT TTT CTC CCT CCA CCA CCC CCC ATG 1390  
 Pro Lys Ser Ala Pro Trp Asn Ser Phe Leu Pro Pro Pro Pro Met  
 185 190 195 200

CCA GGG CCA AGA CTG GGA CCA GGA AAG GTAAACCTTC TATGAAAGTT 1437  
 Pro Gly Pro Arg Leu Gly Pro Gly Lys  
 205

TTCCAGAAAA TAGTTAATGT CGGGACATTT AACCTCTCTG TTAACATAATT TGTAGCTCTC 1497

CCACAAATAT TCTGGGTAAT TATTTTATC CTTTGGTTT TGAGTCCTTT TTATTCCTAT 1557

CATATTGAAA TTGGTAAGTT AATTTTCCTT TGAAATATTC CTTATAG CCA GGT CTA 1613  
 Pro Gly Leu  
 210

AAA TTC AAT GGC CCA CCA CCG CCA CCG CCA CCA CCA CCA CCC CAC TTA 1661  
 Lys Phe Asn Gly Pro Pro Pro Pro Pro Pro Pro Pro Pro Pro His Leu  
 215 220 225

CTA TCA TGC TGG CTG CCT CCA TTT CCT TCT GGA CCA CCA GTAAGTAAAA 1710  
 Leu Ser Cys Trp Leu Pro Pro Phe Pro Ser Gly Pro Pro  
 230 235 240

AAGAGTATAG GTTAGATTTT GCTTTCACAT ACAATTTGAT AATTACCAGA CTTTACTTTT 1770

TGTTTACTGG ATATAAACAA TATCTTTTTC TGTCTCCAG ATA ATT CCC CCA CCA 1824  
 Ile Ile Pro Pro Pro  
 245

CCT CCC ATA TGT CCA GAT TCT CTT GAT GAT GCT GAT GCT TTG GGA AGT 1872  
 Pro Pro Ile Cys Pro Asp Ser Leu Asp Asp Ala Asp Ala Leu Gly Ser  
 250 255 260

ATG TTA ATT TCA TGG TAC ATG AGT GGC TAT CAT ACT GGC TAT TAT ATG 1920  
 Met Leu Ile Ser Trp Tyr Met Ser Gly Tyr His Thr Gly Tyr Tyr Met  
 265 270 275

GTAAGTAATC ACTCAGCATC TTTTCTGAC AATTTTTTTG TAGTTATGTG ACTTTGTTTG 1980

GTAAATTTAT AAAATACTAC TTGAACTGCA GCCTAATAAT TGTTTTCTTT GGGATAACTT 2040

TTAAAGTACA TTAAAAGACT ATCAACTTAA TTTCTGATCA TATTTTGGTG AATAAAATAA	2100
GTAAATGTC TTGTGAAACA AAATGCTTTT TAACATCCAT ATAAAGCTAT CTATATATAG	2160
CTATCTATGT CTATATAGCT ATTTTTTTTA ACTTCCTTTT ATTTTCCTTA CAG GGT Gly	2216
TTC AGA CAA AAT CAA AAA GAA GGA AGG TGC TCA CAT TCC TTA AAT Phe Arg Gln Asn Gln Lys Glu Gly Arg Cys Ser His Ser Leu Asn 280 285 290	2261
TAAGGAGTAA GTCTGCCAGC ATTATGAAAG TGAATCTTAC TTTTGTAAAA CTTTATGGTT	2321
TGTGGAAAAC AAATGTTTTT GAACAGTTAA AAAGTTCAGA TGTTAAAAAG TTGAAAGGTT	2381
AATGTAAAAC AATCAATATT AAAGAATTTT GATGCCAAAA CTATTAGATA AAAGGTTAAT	2441
CTACATCCCT ACTAGAATTC TCATACTTAA CTGGTTGGTT ATGTGGAAGA AACATACTTT	2501
CACAATAAAG AGCTTTAGGA TATGATGCCA TTTTATATCA CTAGTAGGCA GACCAGCAGA	2561
CTTTTTTTTA TTGTGATATG GGATAACCTA GGCATACTGC ACTGTACACT CTGACATATG	2621
AAGTGCTCTA GTCAAGTTTA ACTGGTGTCC ACAGAGGACA TGGTTTAACT GGAATTCGTC	2681
AAGCCTCTGG TTCTAATTTT TCATTTGCAG GAAATGCTGG CATAGAGCAG CACTAAATGA	2741
CACCACTAAA GAAACGATCA GACAGATCTG GAATGTGAAG CGTTATAGAA GATAACTGGC	2801
CTCATTTCTT CAAAATATCA AGTGTTGGGA AAGAAAAAAG GAAGTGAAT GGGTAACTCT	2861
TCTTGATTAA AAGTTATGTA ATAACCAAAT GCAATGTGAA ATATTTTACT GGACTCTTTT	2921
GAAAAACCAT CTAGTAAAAG ACTGGGGTGG GGGTGGGAGG CCAGCACGGT GGTGAGGCAG	2981
TTGAGAAAAT TTGAATGTGG ATTAGATTTT GAATGATATT GGATAATTAT TGGTAATTTT	3041
ATGGCCTGTG AGAAGGGTGT TGTAGTTTAT AAAAGACTGT CTTAATTGTC ATACTTAAGC	3101
ATTTAGGAAT GAAGTGTTAG AGTGTCTTAA AATGTTTCAA ATGGTTTAAC AAAATGTATG	3161
TGAGGCGTAT GTGGCAAAT GTTACAGAAT CTAAGTGGTG GACATGGCTG TTCATTGTAC	3221
TGTTTTTTTC TATCTTCTAT ATGTTTAAAA GTATATAATA AAAATATTTA	3271

## (2) INFORMATION FOR SEQ ID NO:23:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 637 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GATCTGCCTT CCTTCCTGCC CCCATGTTTG TCTTTCCTTG TTTGTCTTTA TATAGATCAA	60
GCAGGTTTTA AATTCCTAGT AGGAGCTTAC ATTTACTTTT CCAAGGGGGA GGGGGAATAA	120
ATATCTACAC ACACACACAC ACACACACCA CACTGGAGTT CGAGACGAGG CCTAAGCAAC	180
ATGCCGAAAC CCCGTCTCTA CTAAATACAA AAAATAGCTG AGCTTGGTGG CGCACGCCTA	240
TAGTCCTAGC TACTGGGGAG GCTGAGGTGG GAGGATCGCT TGAGCCCAAG AAGTCGAGGC	300
TGCAGTGAGC CGAGATCGCG CCGCTGCACT CCAGCCTGAG CGACAGGGCG AGGCTCTGTC	360
TCAAAACAAA CAAACAAAAA AAAAAAGGAA AGGAAATATA ACACAGTGAA ATGAAAGGAT	420
TGAGAGAAAT GAAAAATATA CACGCCACAA ATGTGGGAGG GCGATAACCA CTCGTAGAAA	480
GCGTGAGAAG TTACTACAAG CGGTCCTCCC GGGCACCGTA CTGTTCCGCT CCCAGAAGCC	540
CCGGGCGCCG GAAGTCGTCA CTCTTAAGAA GGGACGGGGC CCCACGCTGC GCACCCGCGG	600
GTTTGCTATG GCGATGAGCA GCGGCGGCAG TGGTGGC	637

## (2) INFORMATION FOR SEQ ID NO:24:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

AGGGCGAGGC TCTGTCTCA

19

## (2) INFORMATION FOR SEQ ID NO:25:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CGGGAGGACC GCTTGTAGT

19

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GCCGGAAGTC GTCACTCTT

19

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GGGTGCTGAG AGCGCTAATA

20

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

TGTGTGGATT AAGATGACTC

20

## (2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CACTTTATCG TATGTTATC

19

## (2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CTGTGCACCA CCCTGTAACA TG

22

## (2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

AAGGACTAAT GAGACATCC

19

## (2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CGAGATGATA GTTGCCCTC

20

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 24 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

AGCTACTTCA CAGATTGGGG AAAG

24

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 20 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

CTCATCTAGT CTCTGCTTCC

20

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 24 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

TGGATATGGA AATAGAGAGG GAGC

24

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 22 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

CACCCTTATA AAAAAACCT GC

22

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 23 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

GAGAAAGGAG TTCCATGGAG CAG

23

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 20 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO



(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

GAGAGGTAA ATGTCCCGAC

20

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

GTGAGAACTC CAGGTCTCCT GG

22

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

TGAGTCTGTT TGACTTCAGG

20

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

GAAGGAAATG GAGGCAGCCA GC

22

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

TTTCTACCCA TTAGAATCTG G

21

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

CCCCACTTAC TATCATGCTG GCTG

24

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

CCAGACTTTA CTTTTTGTTT ACTG

24

(2) INFORMATION FOR SEQ ID NO:45:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

ATAGCCACTC ATGTACCATG A

21

(2) INFORMATION FOR SEQ ID NO:46:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

AAGAGTAATT TAAGCCTCAG ACAG

24

(2) INFORMATION FOR SEQ ID NO:47:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

CTCCCATATG TCCAGATTCT CTTG

24

(2) INFORMATION FOR SEQ ID NO:48:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

AGACTATCAA CTTAATTTCT GATCA

25

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 24 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

TAAGGAATGT GAGCACCTTC CTTC

24

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 25 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

AGACTATCAA CTTAATTTCT GATCA

25

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 24 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

GTAAGATTCA CTTTCATAAT GCTG

24

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

CTTTATGGTT TGTGGAAAAC A

21

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

GGCATCATAT CCTAAAGCTC

20

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

GTAATAACCA AATGCAATGT GAA

23

(2) INFORMATION FOR SEQ ID NO:55:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

CTACAACACC CTTCTCACAG

20

(2) INFORMATION FOR SEQ ID NO:56:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

GGTGTCCACA GAGGACATGG

20

(2) INFORMATION FOR SEQ ID NO:57:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

AAGAGTTAAC CCATTCCAGC TTCC

24

## (2) INFORMATION FOR SEQ ID NO:58:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

Met Ala Met Ser Ser Gly Gly Ser Gly Gly Gly Val Pro Glu Gln Glu  
 1 5 10 15

Asp Ser Val Leu Phe Arg Arg Gly Thr Gly  
 20 25

## (2) INFORMATION FOR SEQ ID NO:59:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

Ser Asp Asp Ser Asp Ile Trp Asp Asp Thr Ala Leu Ile Lys Ala Tyr  
 1 5 10 15

Asp Lys Ala Val Ala Ser Phe Lys  
 20

## (2) INFORMATION FOR SEQ ID NO:60:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

His Ala Leu Lys Asn Gly Asp Ile Cys Glu Thr Ser Gly Lys Pro Lys  
 1 5 10 15

Thr Thr Pro Lys Arg Lys Pro Ala Lys Lys Asn Lys Ser Gln Lys Lys  
 20 25 30

Asn Thr Ala Ala Ser Leu Gln Gln  
 35 40

## (2) INFORMATION FOR SEQ ID NO:61:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 67 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

```

Trp Lys Val Gly Asp Lys Cys Ser Ala Ile Trp Ser Glu Asp Gly Cys
 1             5             10             15
Ile Tyr Pro Ala Thr Ile Ala Ser Ile Asp Phe Lys Arg Glu Thr Cys
                20             25             30
Val Val Val Tyr Thr Gly Tyr Gly Asn Arg Glu Glu Gln Asn Leu Ser
    35             40             45
Asp Leu Leu Ser Pro Ile Cys Glu Val Ala Asn Asn Ile Glu Gln Asn
    50             55             60
Ala Gln Glu
    65

```

## (2) INFORMATION FOR SEQ ID NO:62:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 51 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

```

Asn Glu Asn Glu Ser Gln Val Ser Thr Asp Glu Ser Glu Asn Ser Arg
 1             5             10             15
Ser Pro Gly Asn Lys Ser Asp Asn Ile Lys Pro Lys Ser Ala Pro Trp
    20             25             30
Asn Ser Phe Leu Pro Pro Pro Pro Pro Met Pro Gly Pro Arg Leu Gly
    35             40             45
Pro Gly Lys
    50

```

## (2) INFORMATION FOR SEQ ID NO:63:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:



Pro Gly Leu Lys Phe Asn Gly Pro Pro Pro Pro Pro Pro Pro Pro  
 1 5 10 15

Pro His Leu Leu Ser Cys Trp Leu Pro Pro Phe Pro Ser Gly Pro Pro  
 20 25 30

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

Ile Ile Pro Pro Pro Pro Pro Ile Cys Pro Asp Ser Leu Asp Asp Ala  
 1 5 10 15

Asp Ala Leu Gly Ser Met Leu Ile Ser Trp Tyr Met Ser Gly Tyr His  
 20 25 30

Thr Gly Tyr Tyr Met  
 35

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

Gly Phe Arg Gln Asn Gln Lys Glu Gly Arg Cys Ser His Ser Leu Asn  
 1 5 10 15

50

85